

DOCUMENT RESUME

ED 147 341

TM 006 653

TITLE A Process for the Evaluation of Training.
INSTITUTION Civil Service Commission, Washington, D.C. Training Leadership Div.
PUB DATE [76]
NOTE 36p.; Best copy available
EDRS PRICE MF-\$0.83 HC-\$2.06 Plus Postage.
DESCRIPTORS *Educational Assessment; *Educational Programs; *Evaluation Methods; Evaluation Needs; Federal Government; *Government Employees; Guidelines; *Job Training; Manpower Development; *Program Evaluation; Trainees; Training Objectives

ABSTRACT

The process of evaluating the training of federal civil service employees is summarized in three phases: (1) the focus--establishing the scope and objectives of the evaluation effort, (2) the plan--developing a blueprint for conducting the project, and (3) the implementation--obtaining necessary data, interpreting them, and providing conclusions and recommendations to the client. This introductory pamphlet includes flow charts, a brief glossary, and an annotated list of 20 references. (Author/CTM)

 * Documents acquired by ERIC include many informal unpublished *
 * materials not available from other sources. ERIC makes every effort *
 * to obtain the best copy available. Nevertheless, items of marginal *
 * reproducibility are often encountered and this affects the quality *
 * of the microfiche and hardcopy reproductions ERIC makes available *
 * via the ERIC Document Reproduction Service (EDRS). EDRS is not *
 * responsible for the quality of the original document. Reproductions *
 * supplied by EDRS are the best that can be made from the original. *

a training publication by the



U.S. Civil Service Commission
Bureau of Training

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM

ED147341

A Process for the Evaluation of Training



TM006 653

Training Leadership Division

BEST COPY AVAILABLE

A

Process for the Evaluation of Training_____●

**U.S. Civil Service Commission
Bureau of Training
Training Leadership Division**

Foreword

This pamphlet is the first in a series of U.S. Civil Service Commission publications on evaluation of training. It is presented to:

Broaden the concept of evaluation among trainers and managers;

Specify a sequence for conducting evaluations;

Foster the development of a common vocabulary of evaluation among trainers;

Assist trainers in the self-assessment of their evaluation knowledge and skills.

Intentionally brief, this introductory work is intended to familiarize the reader with the evaluation process and provide a framework for structuring it. "How-to" guides providing step-by-step guidance on the performance of specific evaluation activities will subsequently be issued by the Commission as companion pieces to this paper. Taken together, they will provide a comprehensive guide to the evaluation process. ●

Table of Contents_____●

Introduction	ii
Glossary of terms	iv
Flowchart of the Process	viii
Overview	1
Phase I: Focus	3
Phase II: Plan	9
Phase III: Implement	15
Summary	20
Annotated Bibliography	21

Introduction

Why evaluate? When is evaluation warranted? What resources are required to conduct an evaluation? What information needs to be gathered in the course of an evaluation? If these questions reflect some of your concerns about evaluation, then this pamphlet is for you. It describes a process of activities and decisions necessary to plan and conduct evaluation of most types of training programs.

Explicitly or implicitly, we evaluate—use information to assign value or determine worth—every day, be it when we purchase goods and services or when we make decisions concerning the activities we pursue. Usually these evaluations consist of rapid mental processes based on little information; seldom are they planned or systematic. The definition of evaluation employed herein, however, objectifies and makes explicit this process:

Evaluation is a planned process which provides specific reliable information about a selected topic, problem, or question for purposes of determining value and/or making decisions.

To produce information which will be useful for decision-making, an evaluation effort must be carefully conceived, planned, and carried out.

The process for evaluating training described in this pamphlet provides a systematic approach for formulating and implementing an overall evaluation plan. While the steps are depicted in flowchart form, the process at the same time has substantial flexibility. It allows the evaluator freedom in the choice of methodology, design, and instrumentation, as well as in the modes of interaction with relevant parties. Consequently, the process can accommodate a wide range of interests, timetables, and budgets. Here, we are addressing the problem of evaluating employee training. Therefore, the examples and applications in the pamphlet relate specifically to training situations.

Both managers and those actually involved in conducting evaluations will find the process useful. The process provides a tested sequence for obtaining information to make decisions. While managers are not likely to carry out evaluations themselves, they are often responsible for making yes/no decisions about training, so they need to know what an evaluation process involves. This brief description can meet that need.

The more obvious users of the process are the individuals directly involved in conducting evaluations. However, the matter of training evaluation is of concern to any trainer, supervisor, or manager intent on doing his/her job better. While some organizations are fortunate enough to have individuals who are full-time evaluators, usually the task is assigned to a person who has other responsibilities. The process is applicable in either situation.

Whether the purpose of the evaluation is to make the training the best it can be, or to make a go/no go decision about implementing or keeping it, use of the process requires that there be a decision-maker in need of information about the training. The decision-maker can be the evaluator's manager, the evaluator or trainer, or other individuals within or external to the organization.*

This pamphlet will emphasize the evaluation of **completed** training programs, because, at present, the call for evaluation is more often directed to completed training programs than it is to the evaluation of programs as they are being developed. However, the process is also applicable to those situations where evaluation is a conscious activity during the developmental stages of a training program. The experience of the evaluator will dictate the extent to which the process is adapted to new situations. As an individual's skill increases, the process can be applied to an ever-increasing variety of programs. ●

*Throughout this paper, the individual decision-maker who needs the information will be referred to as client.

Glossary

- Behavior** An assessment of individual performance on the job. (See **Data**.)
- Client** The person(s) or organization(s) for which an evaluation is conducted.
- Control group** A group of persons who do not participate in the training program but who are similar in all relevant respects to those who do participate. Control groups are used as a basis for comparison.
- Correlation** A statistical process which shows the degree to which two or more events or objects are related to each other. In evaluations of training program effectiveness, correlations may be calculated to determine the relationship among factors affecting training results.
- Data** Factual material from which conclusions may be drawn. When evaluating training, four categories of data are commonly obtained: **Reaction, Learning, Behavior, and Results**.
- Data analysis** The process of finding meaning in data, thereby producing information.
- Data processing** The tabulating, grouping, or coding of data in order to prepare them for analysis.
- Design** A strategy which the evaluator uses to collect data. The design usually specifies who will be measured (**experimental group, control group**), and when they will be measured (**pretest, posttest**). The purpose of the design is to guard against the possibility that something other than the treatment causes the observed effects of the training program.
- Evaluation** A planned process which provides specific reliable information about a selected topic, problem, or question for purposes of determining value and/or making decisions.
- Experimental design** A data collection strategy which attempts to control as many rele-

vant and irrelevant **variables** as possible. An experimental design gains its rigor by using **control groups** and **random** selection and assignment of individuals to groups.

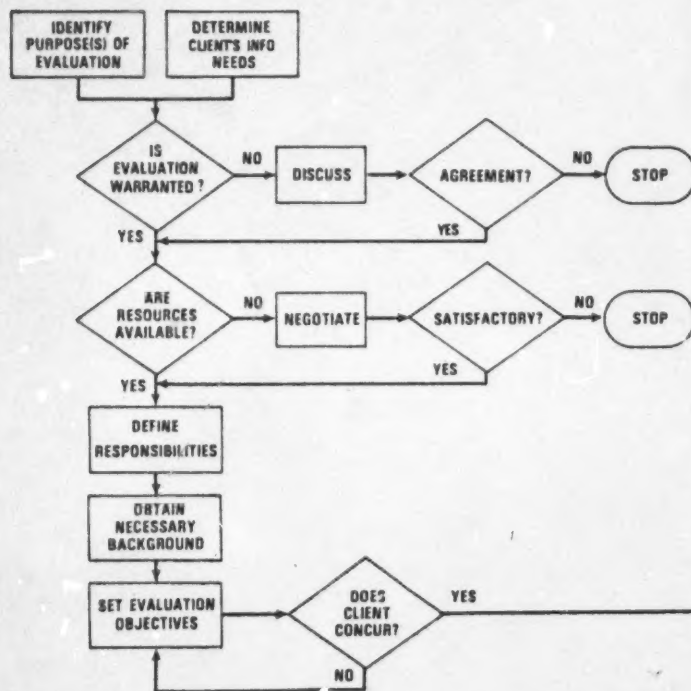
- Experimental group** A group of subjects who receive the experimental **treatment** in a design.
- Learning** The principles, facts, skills, and attitudes that participants gain from training. (See **Data**.)
- Population** The total of any kind of units under study by an evaluator.
- Posttest, Postcourse, or Posttraining measurement** A measurement taken after the training program has ended. The resulting information can be used to determine whether or not trainees have achieved training objectives. If compared to a **pretest** the posttest provides a measure of change possibly attributable to training.
- Pretest, Precourse, or Pretraining measurement** A measurement taken before the training program begins, or during its early stages. The resulting information may be used to provide instructional designers with a picture of the skills and abilities of the average entering trainee, or may be used to give the instructor an idea of how much the group already knows relative to the learning objectives. A pretest also provides a baseline for comparison against a **posttest**.
- Procedures** Instruments, devices, processes used to obtain **data** for evaluation.
- Random assignment or selection** The selection of individuals in such a way that all have an equal probability of being included, and the selection of one individual has no influence on the selection of any other individual.
- Reaction** An indication of how well the trainees liked a particular program including materials, instructors, facilities, methodology, content, etc. (See **Data**.)
- Reliability** The degree to which evaluation procedures measure a given characteristic consistently.
- Results** The impact of training on the organization or job environment. (See **Data**.)

- Rigor A term used to describe the amount of control exerted by a design and the consequent precision that can be attributed to the findings. The more rigorous the design, the greater is the level of confidence that one can have in the findings of an evaluation.
- Sample A subset of the population, usually selected to be representative of the whole group being studied.
- Statistic A measure which describes a sample.
- Validity The degree to which a device or instrument (**procedure**) measures what it was intended to measure.
- Variable Something that is apt or likely to change over time. One purpose of a design is to control for (limit the variability of) irrelevant variables so that the effect of relevant variables may be observed. ●

A Process for the Evaluation of Training

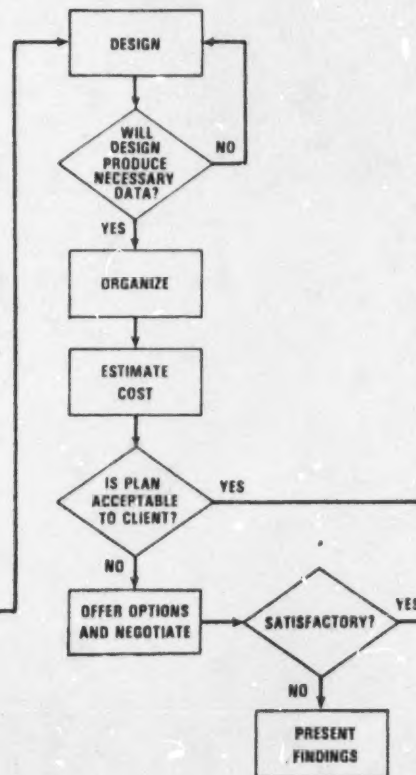
Focus

TO ESTABLISH THE EXTENT
AND OBJECTIVES OF THE EVALUATION EFFORT



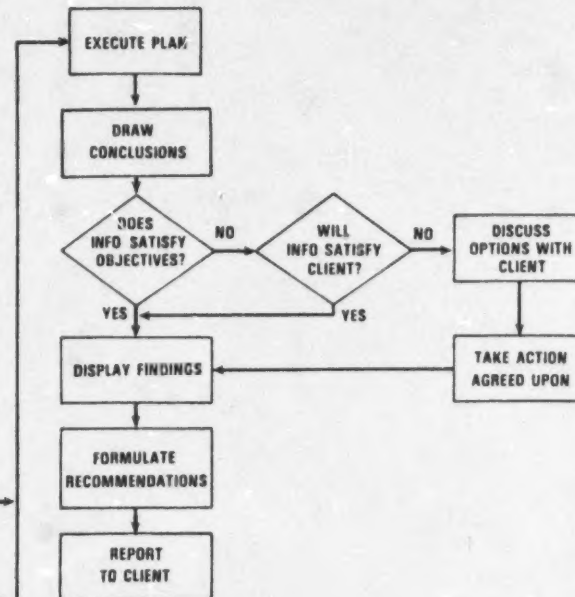
Plan

TO PRODUCE A BLUEPRINT FOR
CONDUCTING THE PROJECT



Implement

TO OBTAIN NECESSARY DATA, INTERPRET THEM,
AND PROVIDE INFO TO THE CLIENT



Overview

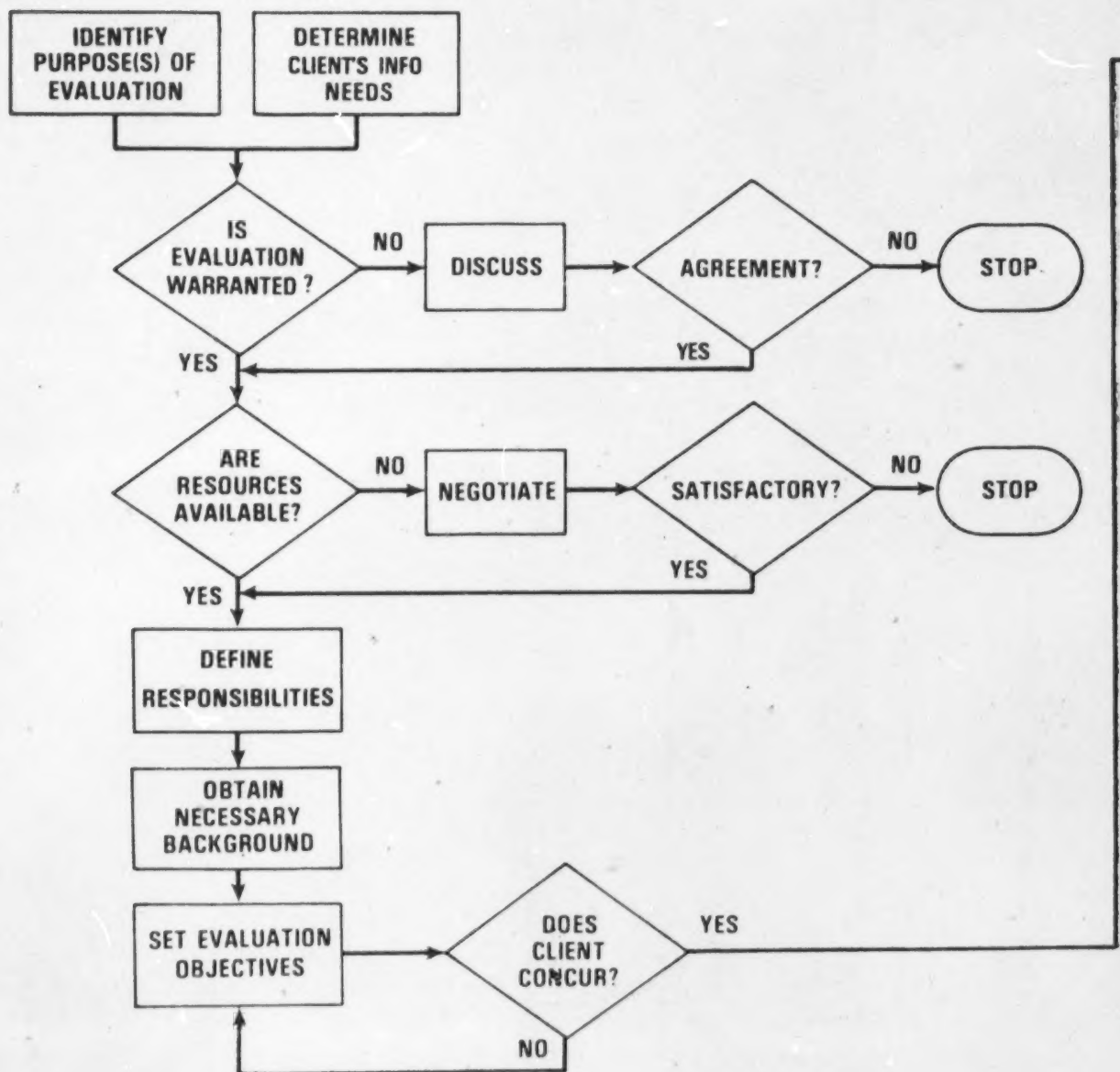
This evaluation process consists of three phases: **FOCUS**, **PLAN**, and **IMPLEMENT**. **FOCUS** comprises the work which establishes the general scope of the evaluation effort. It is here that the evaluator studies the training program to be evaluated and works with the client to establish the extent of the evaluation.

PLANNING produces the blueprint for conducting the project. The plan specifies the data to be collected; the sources from which they will be obtained; the times at which measurements will be taken; the ways in which the data will be gathered; and the schedule for accomplishing all necessary tasks.

During the **IMPLEMENTATION** phase the evaluation plan is put into action. Activities include collecting, tabulating, and analyzing data; formulating conclusions and recommendations; and organizing and presenting the results.

FOCUS

To Establish the Extent and Objectives of the Evaluation Effort



Phase I: Focus

The FOCUS phase calls for close interaction with the client of the evaluation to decide: the purposes the evaluation will serve; the questions it will answer; who will make use of the outputs; whether resources will be available in support of the process; and how the evaluator's role is to be defined.

The evaluator begins by exploring the client's reasons for initiating the project: What prompted the client to request an evaluation? The answer may be as simple as fulfilling an agency or division requirement, or the client may have different or more complex motives. Indications that there are problems in the organization or with the training may signal the client to look to evaluation for some answers. Justifying the allocation of training funds to various programs is another motivating reason for evaluation.

Identify Purpose(s) of Evaluation

Whatever the client's stated reasons may be, training evaluations are usually conducted for the purpose of **making decisions** about two major areas:

*improvements to be made in the training program itself; or
action to be taken regarding the future of the program.*

As examples of the latter, decisions are made as to whether the program should be continued and who should attend it. A client external to the training organization needs to decide whether to buy the training or whether to send employees to one or another course. The client deciding whether to make improvements in the training program itself is concerned with how to run the training more efficiently, maximize the effectiveness of the instructors, increase trainee participation, or other areas.

Determine Client's Information Needs

In any case, the evaluation conclusions should supply the client with an information base to make these decisions. Therefore, the evaluator needs to identify specifically the types of information that are necessary or useful to the client in making the decision.

Let us suppose that the client has requested evaluation for the purpose of making improvements in the training. The specific information needed to do so might relate to whether the program's objectives are achieved; the accuracy of the task analysis preceding training development; instructor skills; the effectiveness of training materials; and so on. For the client inter-

ested in whether the program is cost beneficial, however, the relevant information might center more on changes in employee productivity or measurable improvements in organizational operations. Many evaluations will produce information on **both** the classroom processes and the job/organizational effects of the training since the two are so closely intertwined. The important point is that the evaluator identify precisely what type of information is expected from the evaluation. This can mean the difference between an evaluation report that is filed away and ignored, or one that is accepted and utilized as a basis for making decisions.

Inherent in the purposes for evaluating is the question of what will become of the findings. The evaluator may find that the client's purposes for requesting the evaluation are well-justified from the standpoint of organizational pressures and needs; however, they are not always valid reasons for conducting a full-scale evaluation. At this time the evaluator and client must decide together whether the potential outcomes of the evaluation warrant the effort it will require.

To answer the question "**IS EVALUATION WARRANTED?**" a number of factors should be considered: What usefulness will the findings have? Does the project have both the active backing of the client and the endorsement of others with whom the evaluator will be working? Are changes likely to occur as a result of the study? How will the findings be used? Who in addition to the client will receive the report? Who will be affected by the resulting changes? These questions can form the basis for discussion to resolve any doubts that the evaluation is justified.

Sometimes, the reasons for evaluating may not seem entirely adequate. One such situation is when the decision-makers have already determined the future of the training program. If the program is already packaged and distributed, and is not subject to change, then an evaluation can have little or no impact. Similarly, if management has already decided that a particular training program should be mandatory for supervisors, they may not be receptive to findings which indicate otherwise. The organizational and political realities of the situation should be as clear as possible before the evaluation proceeds.

Occasionally, it happens that the client's and evaluator's differing viewpoints cannot be resolved satisfactorily; in those instances the evaluator should weigh alternative courses of action, including continuing on the client's terms or terminating the project. Obviously, it is easier for the external evaluator to call a stop to the effort than for an internal one. However, both need to consider the options in light of their situation and convictions.

Assuming that solid reasons exist for conducting an evaluation, the evaluator can attend to the practical considerations associated with answering: **ARE RESOURCES AVAILABLE?** Even though an evaluation may ultimately pay for itself many times over in reduced costs, the activity necessarily consumes resources.

The client wishes to know what it will take and the evaluator must ascertain whether what it will take will be available. To answer the question, then, a preliminary sizing up of the project takes place. The evaluator asks:

When must the information be provided in order for it to be useful?

What level of certainty is required of the answers?

To what degree will staff be made available to the evaluation?

What and how much of other resources can be made available, e.g., materials, reproduction, postage, data processing, travel?

The answers to these questions will indicate the client's level of commitment to the evaluation, or to put it another way, the answers to these questions will be a measure of the importance to the client of the information to be pursued by the evaluation.

This step will also aid the evaluator in the planning activities, particularly in selecting an appropriate and feasible design and data-gathering procedures. The type and quantity of available resources will inevitably bear on subsequent crucial decisions such as the size of the sample, how the procedures are developed and by whom, and other design matters which are discussed more thoroughly in the PLAN phase. In any case, this preliminary assessment of resources helps the evaluator establish the real boundaries within which to plan the evaluation.

Define Responsibilities

Next the evaluator needs to reach agreement with the client as to what responsibilities will be assumed by each. For instance, it must be decided who will be accountable for:

- Providing or arranging for access to relevant persons such as trainees, subject-matter experts, and others;

- Gathering or providing the data on which the training will be evaluated;

- Administering testing devices or other evaluation procedures;

- Making decisions to stop or alter the evaluation's course;

- Supplying administrative or support activities linked to the project.

The obligations can differ for the internal and external evaluator. For example, the internal evaluator is likely to have an insider's knowledge of organizational factors which may have a direct bearing on the evaluation. The external evaluator is likely to be more objective as well as more credible to both the client and others interested in examining the evaluation. In both cases, however, the evaluators must be sure that they and their clients fully understand what is expected of each in carrying out the process.

Obtain Necessary Background

Having established what is to be evaluated and in how much detail, the evaluator is free to delve more deeply into the background of the training program. Familiarity should be gained in the following areas:

Subject matter

Even though (s)he may be quite expert in matters of the evaluation process, the evaluator is not expected to know the specific details of every program's subject area. To learn more about the topic of the training, the evaluator may have to consult a subject matter expert or do some indepth reading. The nature of the project should indicate how well the subject matter must be mastered.

Course or program objectives

This is one of the first items the evaluator must determine. What is the program intended to accomplish? Why was it developed? What are the indicators of its success?

When the course package comes equipped with well-stated instructional, i.e., behavioral, objectives, the evaluator's task is greatly simplified. Instructional objectives specify precisely what the participant should be able to do upon completing the training* and, therefore, provide a sound basis for measurement. However, when the course objectives are broadly stated as "goals" or "purposes" of training or when no statements of intended outcomes exist, the evaluator's responsibility is extended. In such cases, an important part of the evaluator's work will be to clarify the objectives with instructors, program developers, the client, or others involved with the training. In general, an evaluation should not proceed without precise, measurable statements of expected behavior.

Program development and package

In addition to the information provided by the program objectives, the evaluator may also need to know: (1) What are the participant selection criteria? (2) Was a training needs analysis done? (3) Does the program still address the original objectives? The client, program developers, instructors, or affiliated staff should have answers to these questions. Background information may also be found in the training materials themselves—for instance, in the instructor's guide. Any evaluation devices that have been built into the program—and ideally several will have—should also be examined, including pretests, criterion tests, examples of performance during training, reaction questionnaires.

Program administration

Finally, the evaluator may need to gather details of the program's actual operation. Such background information includes the number of those taking part in a session; the selection and qualification of instructors; the program's length; and the adherence to participant selection criteria, to name a few.

The activities of collecting relevant background information should not consume an excess of time and effort; however, they should give the evaluator a sufficient foundation to pursue the final activity in FOCUS. . .

Set Evaluation Objectives

On the basis of the material gathered, the evaluator sets forth the project's objectives, the actual accomplishments needed to answer the client's questions. Objectives put the client's questions in a form useful for designing and planning the rest of the effort. If, for example, a manager wants to know whether his/her first-line supervisors possess the necessary knowledge to perform their Equal Employment Opportunity (EEO)-related duties, the objective might read: Measure the extent to which participants learned the facts, legislation, regulations, and applications associated with the course objectives. If, however, the manager's question is, "Are supervisors actively implementing the EEO program on the job?", then the objective might be: Ascertain degree to which supervisors exhibit EEO-related behaviors for which they were trained.

*Instructional objectives should not be confused with evaluation objectives, which are discussed fully in the next section.

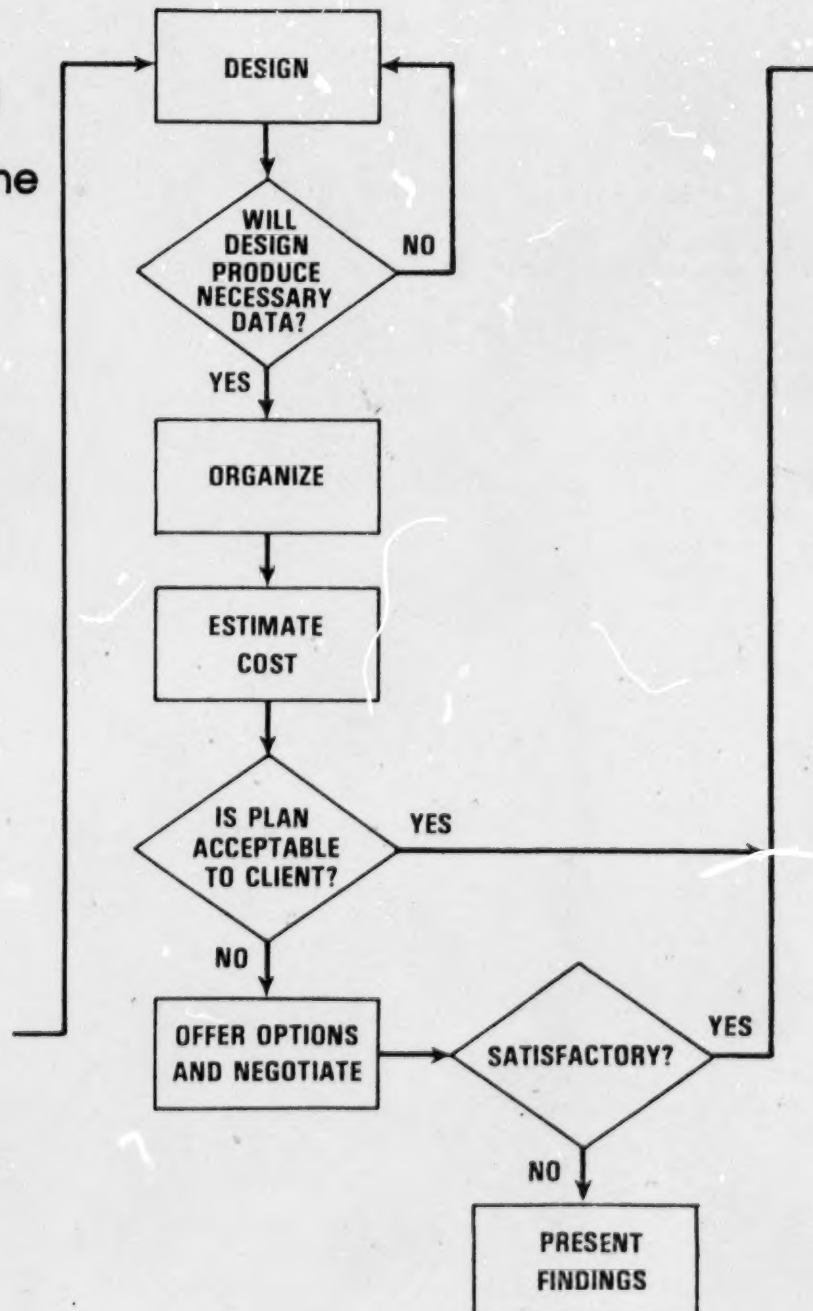
Depending on the client's needs, several objectives may be needed to cover all the necessary ground. In fact, there will probably be a hierarchy of some sort with many lesser objectives to facilitate the achievement of a few broader ones. The ambitiousness of the final objectives will depend on the depth of the undertaking and the nature of the program itself. Developing evaluation objectives may involve substantial time for writing, revising, and consulting with the client in order to produce clear and specific statements.

Assuming that all earlier steps were carried out in close contact with the client, his/her concurrence with the objectives is almost assured. However, on occasion, a number of objectives may have to be revamped to reach full accord. If so, this development presents an opportunity for the evaluator and the client to reach a stronger understanding of how best to continue the process.

In summary, the FOCUS phase should help the evaluator lay a solid foundation for the planning of the effort. As a matter of fact, the evaluator will find it useful to document the FOCUS phase. A written statement of understanding, agreed to by both evaluator and client, should mention among other things, what data the end product will include or provide; what organizational material and/or people will be available to the evaluator; the responsibilities of both client and evaluator; and any reimbursement particulars. Such a statement can prove very helpful later, in resolving questions that arise during PLANNING and IMPLEMENTATION. ●

PLAN

To Produce a
Blueprint for
Conducting the
Project



Phase II: Plan

The activities of the PLANNING phase produce a blueprint for conducting the project. Step by step, the evaluator converts the knowledge and needs gleaned from FOCUS into an action plan. Neglecting these steps invites confusion and waste as the project proceeds; performing them carefully paves the way for securing information that clients will use.

Design

Working out the design of the evaluation study is the first activity required to develop an evaluation plan. The design refines, organizes, and documents the matters discussed during the FOCUS phase. As the framework of the entire project, the design answers these questions:

What data are needed?

Who will provide them?

When will they be obtained?

How will they be collected, i.e., what procedures will be used?

To answer these questions is to begin to address evaluation systematically. This means that the gathering of data is carefully designed in advance and provision is made for analysis. The reward of such work is in the confidence that may be placed in the conclusions of the evaluation. A well-designed evaluation may be expected to produce data from which meaning can be derived and from which accurate conclusions may be drawn.

Where no formal measurements are possible, the knowledgeable judgments of trainers, administrators, and, on occasion, subject matter experts are used, but it is important to be aware of the limitations of informal evaluation.

To answer the four design questions posed above, the evaluator assesses the client's information needs in light of the characteristics of the several designs described in the literature.¹ Then from among the various appropriate evaluation designs, (s)he selects the one (sometimes more than one is desirable) deemed "best" in the particular situation. A "best" design is one that:

Is capable of meeting plausible objections that the findings are due to something other than the training itself;

¹The standard designs range from pre-experimental and quasi-experimental to true experimental. The classic delineation of the range of evaluation designs is authored by Campbell and Stanley (see Bibliography). A simpler treatment may be found in Chabotar's *Evaluation Guidelines for Training Programs* (see Bibliography).

Is suitable to the client in that it will provide for the necessary amount of confidence in the resulting information; and

Is feasible, given both resource and practical constraints.

Balancing all three of these criteria requires, not only knowledge of the major designs, but creative use of all the information at hand. A true experimental design with controls and randomization² will serve the first criterion well, but is rarely feasible in terms of resource and practical constraints. On the other hand, a design that does not provide for any comparisons at all, e.g., one group measured once, would be suitable only to a client whose need for confidence in the information is very low.

Thus, during the early planning stages, the evaluator confronts the matter of the confidence that the client must have in the evaluation's findings. The more confidence a client can have in the conclusions of an evaluation, the more likely they are to be used as the basis of often difficult decisions.

Whatever the rigor of the design, it must pinpoint the:

Categories of data required;

Sources of those data;

Times at which the data will be collected; and

Data-gathering procedures to be used.³

Each of these warrants a closer look.

Data required

Using the evaluation objectives as a guide, specifying the data categories is usually a straightforward matter. Donald Kirkpatrick (1959) has provided four categories useful for describing the types of data that a training evaluation commonly pursues. These are **Reaction, Learning, Behavior, and Organizational Results**.⁴ Reaction and learning are categories of data usually obtained during the training or shortly afterward; behavior and results represent the kinds of data available only in the work environment.

Sources of data

The sources of necessary data may be self-evident from the information gathered earlier about the program or they may need to be ferreted out, or some of each. They may be in the form of documents or people. Where data are subjective, it is advisable not to rely on a single source for an entire category of data. Sources include, for example, trainees, supervisors of trainees, control or competitor groups⁵, training personnel, organizational records, and assessment material associated with previous training or evaluations.

² Randomization is the assignment of members of a population to control and experimental groups in such a way that each has the same probability of being selected. Thus it is a technique which helps prevent bias in the evaluation findings and enhances the evaluator's ability to make generalizations.

³ The gathering of data is here used to include the taking of measurements, whether written, aural, or visual.

⁴ See Glossary under **data** for definitions.

⁵ Competitor groups—any other program participants who are similar and whose results can be compared to those of the program being evaluated.

Joe

Ideally, all members of a chosen data source should be tapped. In reality, however, that is possible only for sources with a very small population. More often than not, evaluators must rely on a portion of an entire population. That portion is known as the **sample**. The data provided by the sample group(s) are then the basis for generalizations about the training. Assessments of programs which train hundreds of employees yearly usually rely heavily on samples of trainees. To do otherwise would be inefficient and unnecessary because sampling is perfectly adequate for most situations.

Timing

The timing of measurements in evaluation is one key element (the other two are controls or comparisons and randomization) in the ability to attribute noted changes to the training and not to other plausible causes. To serve that purpose, it becomes important to take measurements at carefully chosen times. The times at which data sources are tapped will both depend on and contribute to the rigor of the design.

Data can be collected at any number of times relative to the training experience. Measurements made prior to the training can yield data about trainee entry-level knowledge; about learner, supervisor, and trainee expectations, and most importantly, about learner job performance. **Pre-training measurements** can be made just once or at intervals to trace any fluctuations in performance prior to the program.⁶

Measures taken **during and at the end of training** will reveal whether the training program's immediate objectives were achieved or expectations met. These measurements can be compared to those taken prior to the program. On-the-job measurements **following the training** will yield data about the relatedness of the job behavior of trained personnel to the training received or about the impact of the program on the organization. As in the case of pre-training data, the post-training measurements may be made once or at intervals to detect effects of other stimuli, trends, retention of competencies, etc.

Data-gathering procedures

Procedures are the instruments, devices, or processes used to obtain data. While the questionnaire and the interview are the ones that most often come to mind, the choices available to the evaluator are many and varied.⁷ The data categories indicated by the design serve as a starting point. An interview, for example, may produce insufficient data regarding learning, but may be highly effective in obtaining individual reaction, as well as in gauging on-the-job behavior changes. Before making final selections, it's necessary to weigh the advantages and disadvantages of each procedure deemed suitable for the data category. This will involve considering:

- Requirements of the related evaluation objective(s);
- Accessibility of the data source;
- Practicality in terms of time, skills, and funds required; and
- Accuracy, reliability, and validity of the procedures.

Clearly, determining the suitability of a data collection instrument, device, or process is

⁶This is a reference to a specific design, namely, time series. For information on the uses of this design, see the works cited earlier (Campbell and Stanley, Chabotar, and others in the Bibliography).

⁷"An Aid for the Training Evaluator," useful in selecting procedures, is available from the Office of Agency Assistance and Resource Coordination, Bureau of Training, U.S. Civil Service Commission.

a matter requiring careful consideration of a variety of factors. Two of the most important of these factors—validity and reliability—are discussed later in IMPLEMENTATION.

Formulating the design forces the evaluator to think through the conditions under which the study will take place. In working out the details discussed above, it is easy to lose sight of the original objectives of the evaluation. When the design is completed, then, it is important to step back to consider: **WILL THE DESIGN PRODUCE THE NECESSARY DATA?** Answering this question constitutes two checks:

Will the anticipated data be sufficient to answer the client's questions?

Can the design be expected to produce those data?

This is the evaluator's opportunity to review the objectives to see that the design is indeed complete and properly configured to meet the client's information needs. If the answer is "yes," the business of organizing the study is at hand.⁸

Organize

Organizing the evaluation involves:

Listing, in order, the tasks necessary to carry out the design; and

Calculating times and setting up appropriate schedules.

These steps are sometimes tedious to accomplish, but they are critical for the proper execution of the design and essential to an efficient study.

Laying out tasks

The task list may be as succinct or as detailed as seems appropriate to the situation and the evaluator's experience. However, the more specific the task list, the more accurate subsequent scheduling can be. Whatever the level of detail, the list should include all that needs to be done, from procuring needed external expertise, to setting up groups for testing, to initiating communications with persons who will be involved in the administration of procedures (and possibly training them), to working out how the data will be analyzed and reported.

Scheduling activities

The task list becomes the basis for a master schedule which, again, may be as simple or complex as the evaluation objectives demand. Care should be taken that the measurement times indicated by the design can, in fact, take place accordingly. Estimates of how long tasks will take or how long to allow for questionnaires to be returned, for example, can be crucial. Since some tasks can be more complex than originally thought, estimates should anticipate possible delays and also allow for the slippage which unforeseen problems will cause.

At first glance, the steps involved in organizing an evaluation project may seem unnecessarily detailed. The value of such specification and detail will manifest itself later, however, in time saved and in needed data which would otherwise be lost.

⁸ For aid in judging an evaluation design, see James R. Sanders and Dena H. Mafziger, *A Basis for Determining the Adequacy of Evaluation Designs*, Phi Delta Kappa, Occasional Paper 20, July 1976.

Estimate Cost

Earlier in the process, during the FOCUS phase, the evaluator ascertained whether resources were available to conduct the evaluation. This was a general assessment to verify that the client was willing to invest some amount of time, staff, or funds in the effort. At this stage, the evaluator can form a more precise estimate of the total cost of the study. Each element of the plan can be assessed for time, manpower, travel, data processing, and other costs. With the plan and approximate cost in hand, the evaluator returns to the client with a detailed description of how each objective will be accomplished and at what cost.

The obvious question to be answered next becomes: **IS THE PLAN ACCEPTABLE TO THE CLIENT?** The client's concurrence with the completed plan and its accompanying cost is essential to the conduct of the project. Although the plan may meet with complete acceptance from the client, the evaluator should be prepared to discuss possible modifications.

Options

Modification of objectives

Seeing the full plan, the client may wish to obtain some additional information by expanding the objectives, or, alternately, paring them down to concentrate on major points. The plan would then have to be modified accordingly.

Modification of the design

While it may be advisable to keep the original objectives, the rigor of the design may be open to negotiation. Adjustments in the sample size, in the number of measurements taken, or in the way group members are selected are some possibilities.

Modification of procedures

The evaluator should be ready to propose variations in the chosen procedures, or, to the extent possible, substitution of completely different procedure(s).

Whatever the options being considered, it is the evaluator's responsibility to point out the ramifications of any changes to the client. If the use of a different procedure will produce less reliable data, for example, the client should be apprised of that fact. Similarly, the client should be made aware of any changes in the timetable occasioned by adjustments in the plan.

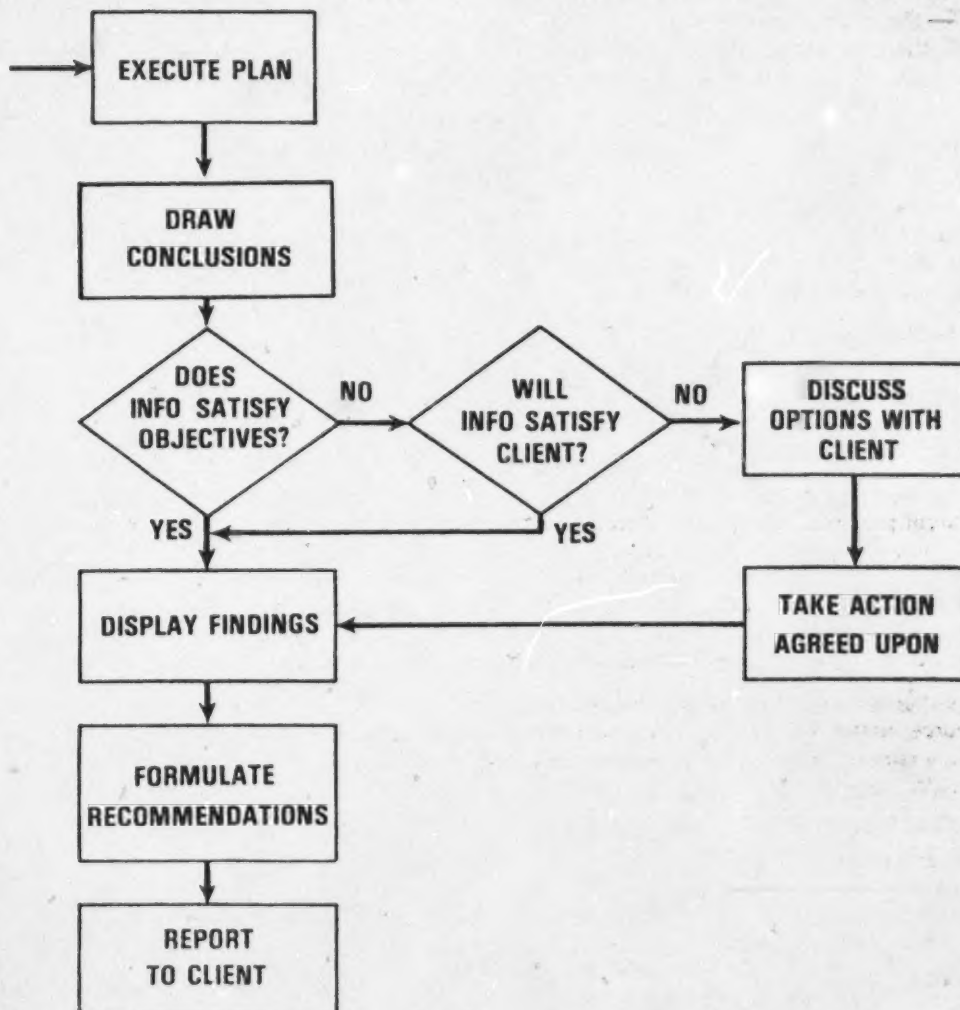
Resource constraints are the usual basis for client problems with an evaluation plan. Hence, if the resource matter was realistically discussed during the FOCUS phase, and if the evaluator has conscientiously adhered to those limits in formulating the plan, there should be few, if any, objections to the plan.

On occasion it may happen that the evaluator's needs will have changed or that resources are no longer available.

If the issues cannot be resolved satisfactorily, the evaluator may have little alternative but to give the client any findings up to this point and terminate the effort. Obviously, the internal evaluator's situation is more delicate, and may not include the luxury of refusing to continue. Nevertheless, the evaluator should point out the undesirability of pursuing an evaluation which cannot yield reliable, usable information. ●

IMPLEMENT

To Obtain Necessary Data, Interpret Them, and Provide Information to the Client



Phase III: Implement _____ ●

Thus far in the process, the evaluator has worked closely with the client, first in focusing the direction of the evaluation effort, then in planning how the study will proceed. To begin the final phase of the process, the evaluator puts the plan into action, using the tasks and schedule developed in PLAN. While this should be a fairly straightforward activity, particularly if the plan has been carefully prepared, unanticipated difficulties may arise at various stages. These may range from uncooperative interviewees, to problems in obtaining sample populations, to complications in data analysis. To the extent possible, the schedule inherent in the plan should be followed. However, when unexpected hurdles appear, the evaluator should keep the client informed of any resulting delays in completion of the project.

Execute Plan

The activities involved in carrying out the plan include development and validation of procedures, collecting, processing, and analyzing data.

Developing and validating procedures

The tasks involved in developing procedures revolve around the complexity of the procedures and their availability in ready-made form. Some points to be considered:

If procedures are being developed in-house, the work should be organized on the basis of when they will be used and the time needed to develop them. Alternate forms of the same device can be developed for testing. This will aid in producing a format which best elicits the desired data.

If instruments from another source are being used, they should be reviewed to ensure that they elicit the data anticipated. Successes, problems, limitations, and results from previous usage can be checked as well. It may be necessary to adapt them to suit the particular training program.

If development of the instrumentation is contracted to another party, the evaluator's responsibility is to outline to the contractor what is desired and allow time for consultation and revision.

Whether the instruments are developed internally or externally, they should be tested on a population comparable to that which will be providing the actual feedback. This helps assure that they are valid, i.e., that they do the job for which they are used. Valid instruments are crucial if the evaluator is to be sure that the right things are being measured. And, if the wrong things are being measured what confidence can the client have in decisions made on the basis of the information? At best, if the instruments do not

measure what they purport to measure, the evaluation may be unusable. At worst, faulty or erroneous conclusions may be drawn which could lead to poor or wrong decisions. These criteria may be used to assess validity: Are the items phrased clearly and precisely? Do they elicit responses yielding the desired data? Is all necessary (and only necessary) data covered?

The **reliability** of the instruments should also be checked during testing to see that the instruments produce the same kinds of data each time they are used, i.e., that the data produced are accurate and consistent. There are a number of variables which may result in inconsistencies among measurements, e.g., variations in the situation or environment, differences in the administration of the instruments, or changes over time in the respondents or subjects. Reliability involves some determination of the extent to which these variables are likely to produce inconsistent measurements.

The results of testing for validity and reliability should be channeled into revisions and subsequent retesting as many times as necessary.

Collecting and processing data

At the times specified in the schedule, the instruments should be administered and the data processed. Conducting interviews, administering questionnaires, observing, and testing are all activities that take place during this period. Data collected from these procedures will rarely be usable by themselves; they must be processed, categorized, coded, and tabulated. This may be done via simple manual tabulation or through complex, computerized processing. The processing of the data puts them in a form which can serve as the basis for analysis.

Analyzing data

Analyzing data is the process of extracting meaning from them. The analysis may take many forms, ranging from a simple review of the data to complex statistical testing. It may be in numerical or narrative form. The nature of the analysis will have been established in the PLAN phase, in accordance with the objectives and the design. At this stage, the evaluator is walking through the steps outlined earlier. Whatever the extent of analysis, the purpose is always to enable the evaluator to draw inferences and conclusions from the data.

Full explanations of how to calculate and use statistics may be found in various of the works cited in the bibliography. Usually, analysis of the data will involve:

Measures of central tendency—the mean, median, or mode.

Measures of dispersion or distribution—such as the range or standard deviation and variance.

Comparisons and relationships—correlations, factor analysis, analysis of variance, ratio tests.

For summarizing a set of scores from one group, a measure of central tendency and spread of the scores is usually used. To compare the scores obtained by two or more groups on a single measure, some correlation or other comparison is calculated.

Draw Conclusions

The evaluation conclusions draw together the various pieces of information and supply the answers to the client's questions. Conclusions should address all of the concerns raised by the client early in the FOCUS activities, as well as point out areas where the data are inconclusive. Important points should be emphasized and care taken to state what conclusions must be "tentatively" rather than "firmly" held.

The evaluator may find it helpful to draw conclusions for each evaluation objective and discuss the conclusions in conjunction with the findings for the appropriate objective. This will aid the user of the report in understanding how the conclusions were reached.

Does Information Satisfy Objectives?

Once the conclusions have been stated, the resulting information can be compared against that called for by the evaluation objectives. This questioning constitutes the evaluator's final check on all the preceding work. If the analysis of the data has, in fact, produced answers to the client's questions, the evaluator can proceed confidently to the recommendations and final reporting.

It is possible, however, that unforeseen circumstances may occasion a discrepancy between the actual evaluation results and the information needed to satisfy the objectives. In that event, the evaluator should consider the question: **WILL THE INFORMATION SATISFY THE CLIENT?** If the evaluation plan has been carefully laid out and implemented, the results should be fairly complete and entirely satisfactory to the client. If, however, substantial information is missing or inaccurate, the evaluator and client can discuss alternative courses of action. The evaluator may find it necessary to submit some estimates of time, personnel, and other costs for another plan of action.

As with the original evaluation plan, the costs will have to be weighed against benefits to be gained from any additional efforts. The client may be willing to fund further work, or (s)he may be satisfied that the additional information is not worth the cost.

Display Findings

When the evaluator is satisfied that the objectives have been met, (s)he can proceed to prepare the report and present the findings to the client. No matter how sophisticated or simple the data analysis may be, a clear display of the results is mandatory. Regardless of how well the earlier steps may have been carried out, the presentation of findings must be clear in order for the evaluation report to be useful. Clarity also minimizes the possibility of misinterpretation of the data.

The data which are selected for presentation should be those which are most relevant to the person receiving recommendations; displays which facilitate decision-making, incorporation of improvements, or other actions are most desirable. Graphic presentations of data are frequently preferable to tables, and, surely, to prose for conveying an overall picture of data to the reader. Pie charts, bar graphs, and frequency polygons give clear pictorial representations of

trends and relationships. In cases where detailed data must be made available as well, the graphs may be accompanied by tables. The graphs and tables become the basis of the narrative description of the results.

Formulate Recommendations

Whereas the conclusions supply the direct answer to the client's information needs, the recommendations speak to the client's decision-making needs. Therefore, recommendations can be the most influential part of the report.

Whenever possible, recommendations for actions concerning the program should offer alternative courses of action. This heightens the likelihood that recommendations will actually influence the decision-maker. Because the evaluator usually has little power to implement the changes, the success of the evaluation is measured by the degree to which it is used by the client. Thus it is all the more important that the client be given clear information and workable courses of action from which to make decisions.

Furthermore, the recommendations should include suggestions for undertaking the appropriate action. For example, an evaluator might recommend that the client revise the participant selection criteria for a training program, but the recommendation should not stop there. Rather, it should also address the question of *how* the criteria should be revised, e.g., by establishing specific prerequisites or by conducting and adhering to a training needs assessment.

If appropriate, recommendations concerning subsequent evaluations may also be helpful. These would address such questions as: Should some instruments be modified or discarded for subsequent evaluations? Should a different design be attempted in future evaluations? How can evaluation be incorporated into the course as a standard process?

Report to the Client

The final report can be organized in whatever way seems appropriate and desirable to the client. Whether the presentation is oral or written, it should contain a number of sections in addition to the data displays.

It is useful to provide a short summary of major conclusions and recommendations at the beginning of the report.

There should be a statement of the objectives of the evaluation.

The program under consideration should be described.

The evaluation study itself should be described: the design(s) chosen, the procedures used, the data collected, and the results obtained. It is important to discuss the degree to which results are attributable to the training. In this section, the evaluator needs to describe clearly the implications of any constraints which were placed on the effort or any events that may have contaminated the results.

Finally, the evaluator should list conclusions emanating from the evaluation and the recommendations for action.

The style in which the evaluation report is presented is determined by the client or audience who will be using it. Some people prefer to read the report thoroughly first, and then discuss questions with the evaluator. Others prefer an oral briefing or presentation first, after which they receive the written report.

The challenge for the evaluator is to present the findings in the manner to which the client will be most receptive while ensuring that the integrity of the findings remains intact. Whatever format is used, every effort should be made for the evaluator and client to meet and discuss the outcomes of the project. At this time, the evaluator may wish to discuss possible follow-up arrangements, so that (s)he will know whether the evaluation report was actually the basis for changes. This would also be the appropriate time for the client to ask questions and arrange for any help needed in implementing the recommendations. Such follow-up assistance is unusual in the case of external evaluators. Internal evaluators, however, are more likely to participate in carrying out recommendations simply because of their proximity to and knowledge of the organization and situation.

Whether or not the evaluator becomes responsible for assisting in carrying out recommendations, (s)he is responsible for clarifying, explaining, or substantiating any of the content of the report. ●

Summary

To FOCUS, PLAN, and IMPLEMENT an evaluation effort is to approach an often difficult, sometimes delicate job, systematically. That is not to say, however, that each step of the process must or should be carried out precisely as described herein. The process suggests a way of conceptualizing, planning, and doing a complex job as straightforwardly as possible.

This process is not limited to any particular type of organization or specific kind of evaluation. It will serve the homespun ones as well as the sophisticated ones, the small as well as the large, because it is a guide to be adapted as each circumstance requires and the judgment of individuals dictates.

The success of an evaluation is not a matter of how much information is obtained; rather it is a function of the influence that it has on decision-making. Regardless of the limitations imposed by resources or situation, a systematic approach to evaluation will increase the likelihood that the work will be useful and utilized.

Bibliography

The Focus phase of training evaluation requires skills mainly in consultation. Comprehensive texts include:

Blake, Robert R., and Jane Srygley Mouton. **Consultation**. Reading, MA: Addison-Wesley, 1976. 484 pp.

The authors identify five major kinds of intervention, five settings for change, and four focal issues, which they organize in a single conceptual scheme.

Lippitt, Ronald, and Gordon Lippitt. "Consulting Process in Action, I and II," **Training and Development Journal**, 29, Nos. 5 and 6 (May and June, 1975), 48-54, 38-41.

In this two-part article the authors examine the phases of the consultation process, the roles of the consultant, and strategies used for intervention.

Two books with specific orientations which may prove useful in both the Focus and Implementation phases are:

Bolsky, Morris. **Guide to Gathering Information in Face to Face Interviews**. Ramsey, NJ: Ramsey-Wallace, 1967. 69 pp. Paper.

This book provides over 100 practical recommendations covering all phases of the interview situation.

Weiss, Carol H. **Evaluation Research: Methods for Assessing Program Effectiveness**. Englewood Cliffs, NJ: Prentice-Hall, 1972. 160 pp. Paper.

Although devoted to program evaluation, it contains many ideas which transfer readily to the evaluation of training programs. Especially useful are the discussions of the *focus* and *implementation* aspects of evaluation.

The Plan phase calls for knowledge of the range and characteristics of research designs. An important reference for those conducting or using evaluations is:

Campbell, D.T., and J.C. Stanley. **Experimental and Quasi-experimental Designs for Research**. Chicago: Rand McNally, 1966.

This monograph has become a classic for evaluators. The emphasis is on the logic of comparative experimentation. The content is technical but easy to read.

Planning can also require the development of measurement instruments, including tests of various sorts. Two useful works on testing:

Blood, Don F., and William C. Budd. **Educational Measurement and Evaluation**. New York: Harper and Row, 1972. 212 pp. Paper.

This publication deals with the construction and use of both observational and paper-and-pencil tests.

Hively, Wells (Ed.). **Domain-Referenced Testing**. Englewood Cliffs, NJ: Educational Technology Publications, 1974. 150 pp. Paper.

Domain-referenced testing is related to evaluation by several articles in this compilation.

The Plan and Implementation phases of training evaluation require the skills of social science research as they are applied in educational research. The Southwest Regional Laboratory for Educational Research and Development has prepared a set of seven self-instructional sequences in instructional product research. The set is available clothbound, and each sequence is available in pamphlet form. Three of the pamphlets are listed below, and each contains a well-selected list of works that deal with the topic in greater detail. These pamphlets are an excellent place for the newcomer to evaluation to start.

Resta, Paul E., and Robert L. Baker. **Selecting Variables for Educational Research**. New York: American Book Company and Van Nostrand Reinhold, 1972. 52 pp.

This pamphlet describes ways to identify and select variables in formulating a research problem. It also explains how to find related research which has already been reported.

Sullivan, Howard J. **Classifying and Interpreting Educational Research Studies**. New York: American Book Company and Van Nostrand Reinhold, 1972. 30 pp.

The author explains how to classify and interpret studies so that one may plan a study which will yield the needed information.

Wolf, Richard M. **Choosing an Appropriate Statistical Procedure**. New York: American Book Company and Van Nostrand Reinhold, 1972. 30 pp.

Wolf offers a table of statistical techniques classified by type, number, and measurement scale of the variables. It is easily understood by those not trained in statistics. After a statistic has been selected, a standard statistics text should be consulted to determine further requirements of the statistic.

The Plan phase calls for a careful assessment of the evaluation design. The following paper includes a checklist for judging designs:

Sanders, James R., and Dena H. Mafziger. **A Basis for Determining the Adequacy of Evaluation Designs**. Phi Delta Kappa, Occasional Paper 20, July 1976.

A number of more general works cover a whole range of evaluation topics, procedures, and principles. A representative sample of these follows:

Anderson, Scarvia B., et al. **Encyclopedia of Educational Evaluation**. San Francisco, CA: Jossey-Bass, 1975. 515 pp.

This reference work contains about 125 short articles on the major concepts and issues in educational evaluation. A classification of the articles is provided so that one can easily identify related topics.

Bergman, Brian A., and Arthur I. Siegel. **Training Evaluation and Student Achievement Measurement: A Review of the Literature**. Lowry Air Force Base, CO: Dept. of the Air Force, 1972.

This publication selectively reviews the current literature related to training evaluation and student achievement measurement, extending over twenty years. A useful summary of "what is new in training evaluation."

Borich, Gary D. (Ed.). **Evaluating Educational Programs and Products**. Englewood Cliffs, NJ: Educational Technology Publications, 1974. 491 pp.

A handbook of articles that discuss the roles, models, and methods of educational evaluation. It covers establishing perspective, planning the evaluation, and analyzing data.

Chabotar, Kent J., and Lawrence J. Lad. **Evaluation Guidelines for Training Programs**. Lansing, MI: Michigan Department of Civil Service, State of Michigan, 1974. 253 pp. Paper.

This guide covers all aspects of training evaluation from planning to making recommendations; it includes self-instructional exercises.

Hodgkinson, Harold L., Julie Hurst, and Howard Levine. **Improving and Assessing Performance: Evaluation in Higher Education**. Berkeley, CA: Center for Research and Development in Higher Education, University of California, 1975. 242 pp.

This publication has a useful annotated bibliography. It also explains education evaluation techniques that are applicable to training evaluation.

Isaac, Stephen, and William B. Michael. **Handbook in Research and Evaluation: A Collection of Principles, Methods, and Strategies Useful in the Planning, Design, and Evaluation of Studies in Education and the Behavioral Sciences**. San Diego, CA: Robert R. Knapp, 1971. 186 pp.

This comprehensive paperbound text covers planning an evaluation, designing research, constructing tests, using statistics, analyzing data and preparing reports.

Kirkpatrick, Donald L. (Ed.). **Evaluating Training Programs**. Madison, WI: ASTD, 1975. 313 pp. Paper.

A collection of sixty articles reprinted from the **Journal of the American Society for Training and Development**. The articles deal with various aspects of training, mainly in industry.

Popham, W. James (Ed.). **Evaluation in Education: Current Applications**. Berkeley, CA: McCutchan, 1974. 585 pp.

Nine essays by recognized scholars in such fields as data-gathering design, cost analysis, standardized testing, and criterion-referenced measurements. ●

The groundwork for this pamphlet was done by Asia Rial Elsbree and Joanne J. Jorz of the Training Leadership Division staff. The text was written by Christine Howe and Asia Rial Elsbree, with contributions from Linda Armstrong and Grazia Narkus-Kramer. ●